



Kernel Analysis in Nsight Compute

July 8, 2020

An abstract network diagram with green nodes and lines on a dark background. The nodes are represented by small, glowing green circles of varying sizes, and the lines are thin, green, semi-transparent lines connecting the nodes in a complex, web-like pattern. The background is a dark, almost black, gradient with some subtle light effects.

Nsight Developer Tools

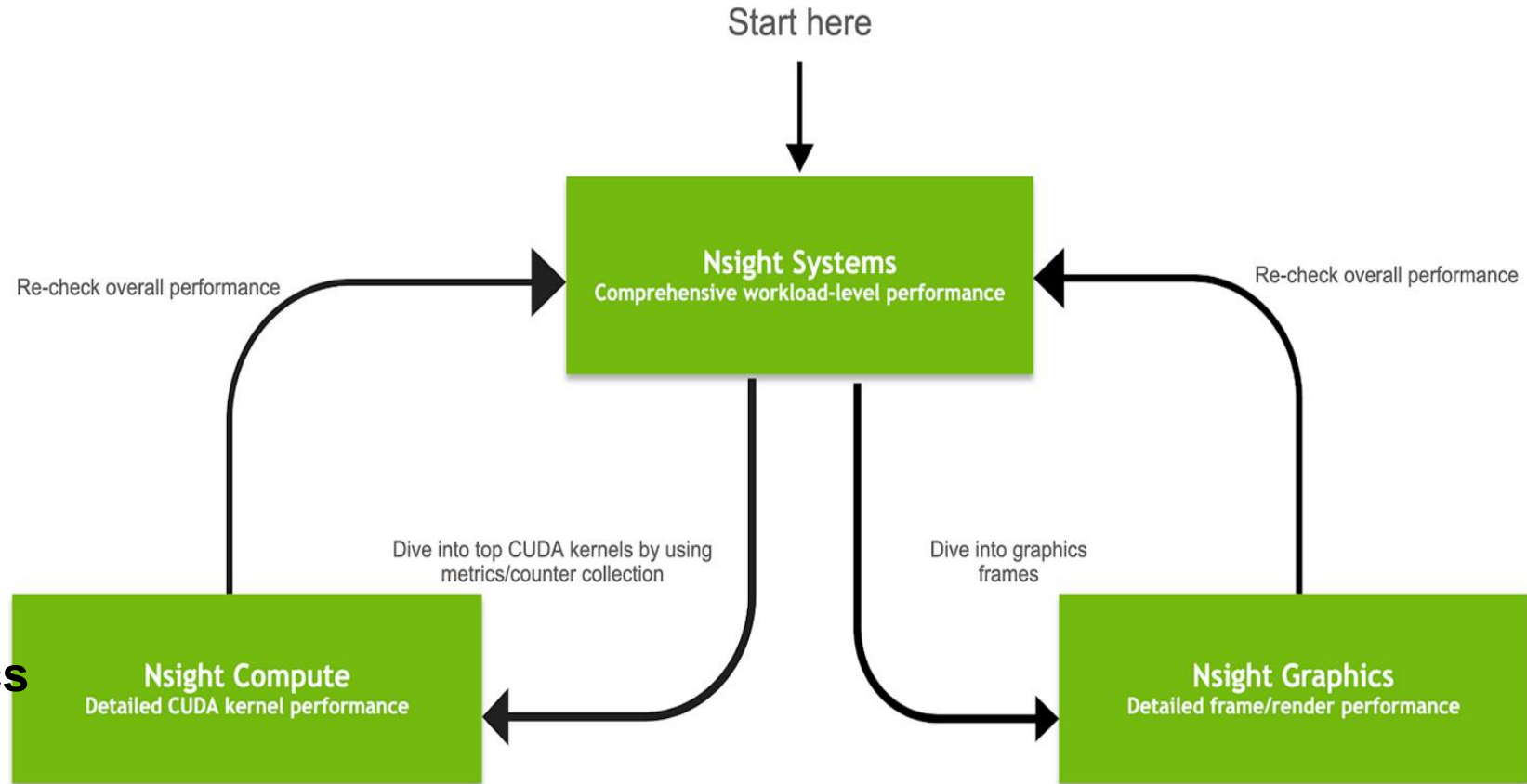
Nsight Product Family

Workflow

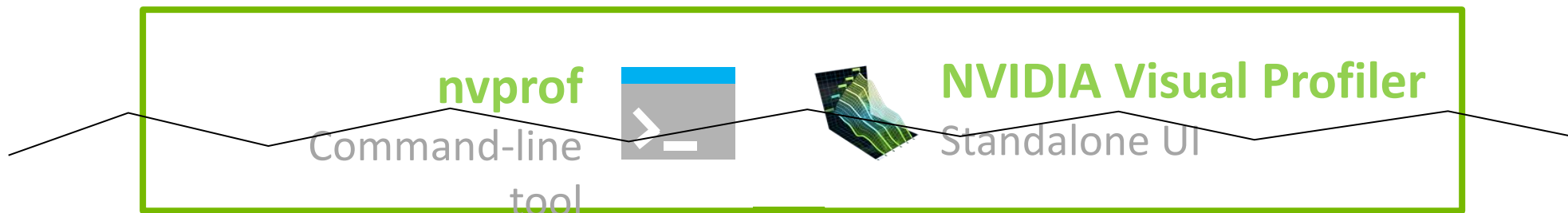
Nsight Systems -
Analyze application
algorithm system-wide

Nsight Compute -
Debug/optimize CUDA
kernel

Nsight Graphics -
Debug/optimize graphics
workloads



Legacy Transition



Nsight Systems

Standalone GUI+CLI

- CPU-GPU interactions & triage
- Low overhead capture
- GPU compute & graphics
- Faster GUI + more data



Nsight Compute

Standalone GUI+CLI

- GPU CUDA kernel analysis & debug
- Very high freq GPU perf counters
- Compare results (diff)
- Incredible statistics & customizable

COLLECTING PROFILES WITH NSIGHT SYSTEMS

```
$ nsys profile --stats=true ./myapp.exe
```

Generated file: `report.qdrep`

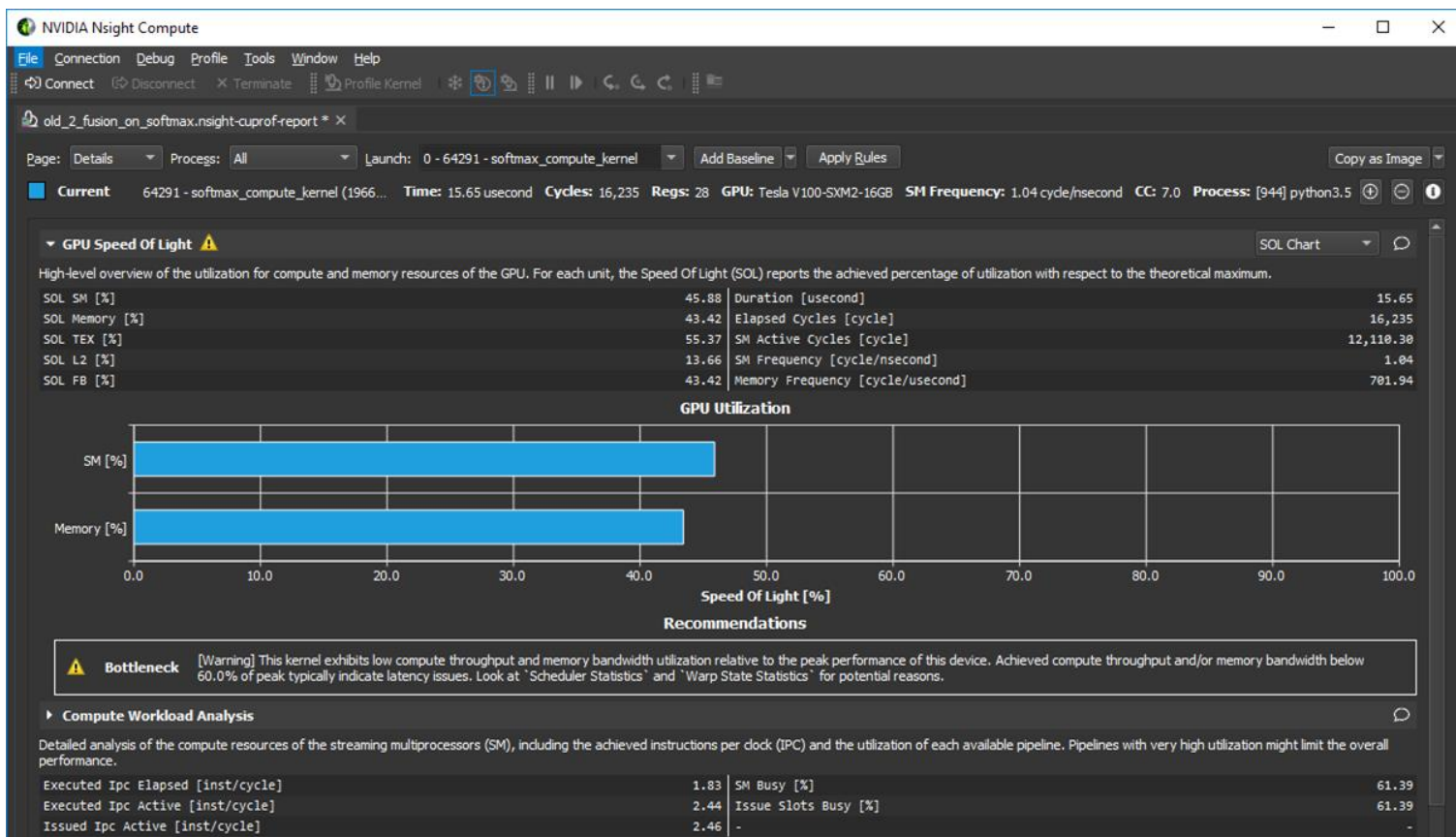
Import for viewing into the Nsight Systems UI

The Nsight Systems UI can also be used for interactive system profiling

An abstract visualization of a network or data structure. It features numerous small, bright green circular nodes scattered across a dark, almost black background. These nodes are interconnected by a dense web of thin, light green lines, creating a complex, web-like pattern. Some nodes are slightly larger or more prominent than others. The overall effect is one of dynamic connectivity and data flow.

Nsight Compute

Nsight Compute



CUDA Kernel profiler

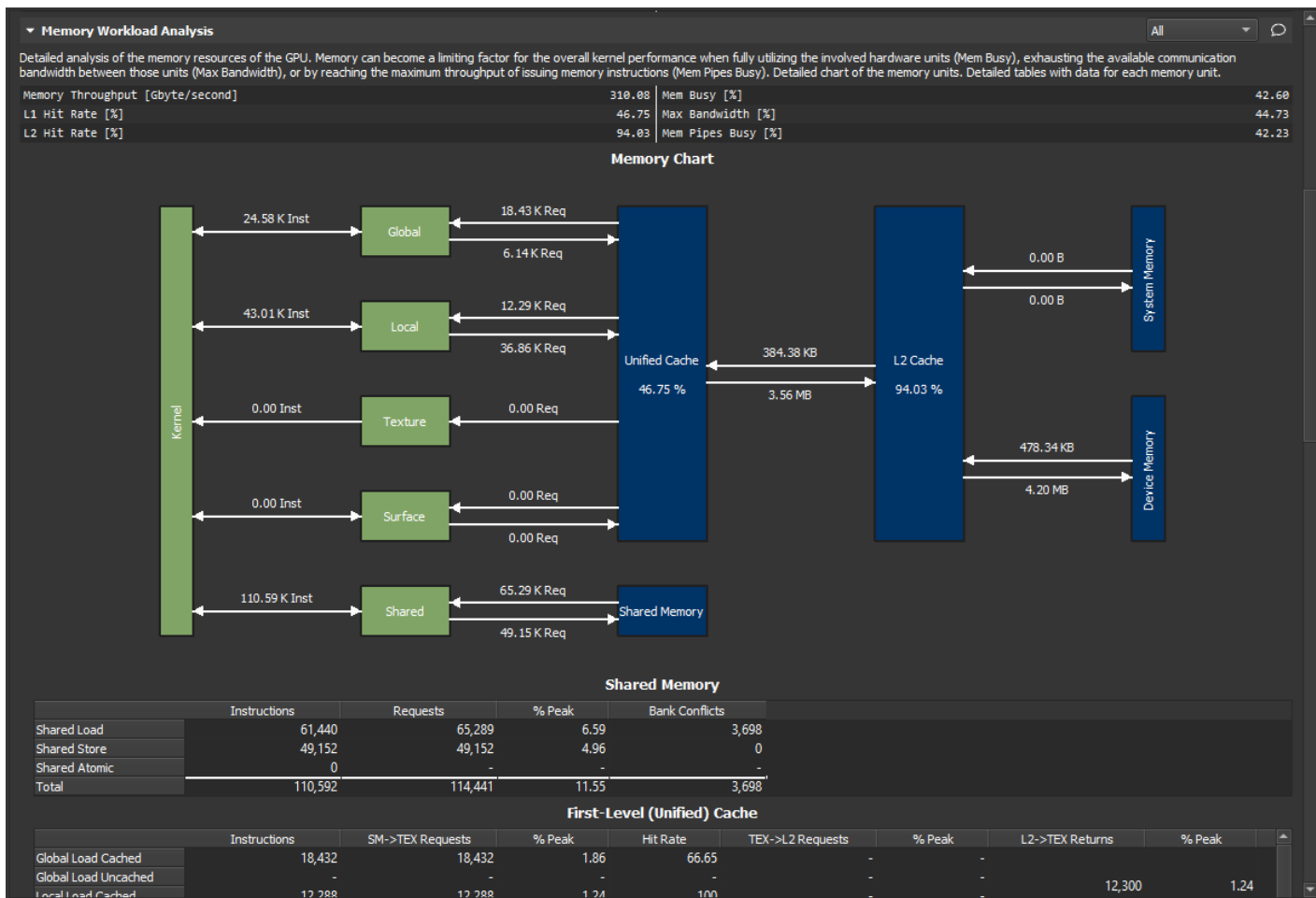
Targeted metric sections for various performance aspects

Customizable data collection and presentation (tables, charts, ...)

UI and Command Line

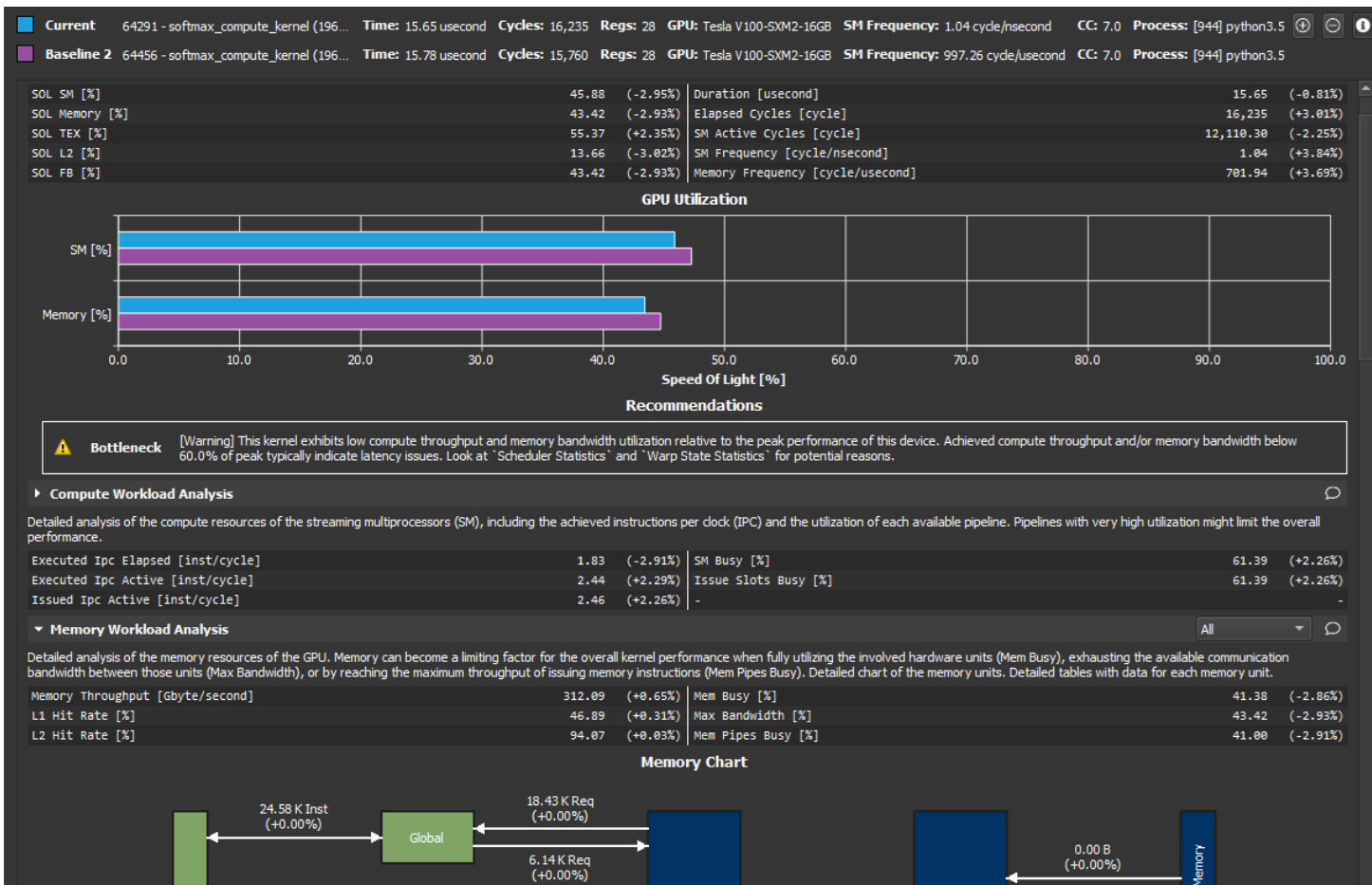
Python-based rules for guided analysis (or post-processing)

Nsight Compute



Detailed memory workload analysis chart and tables

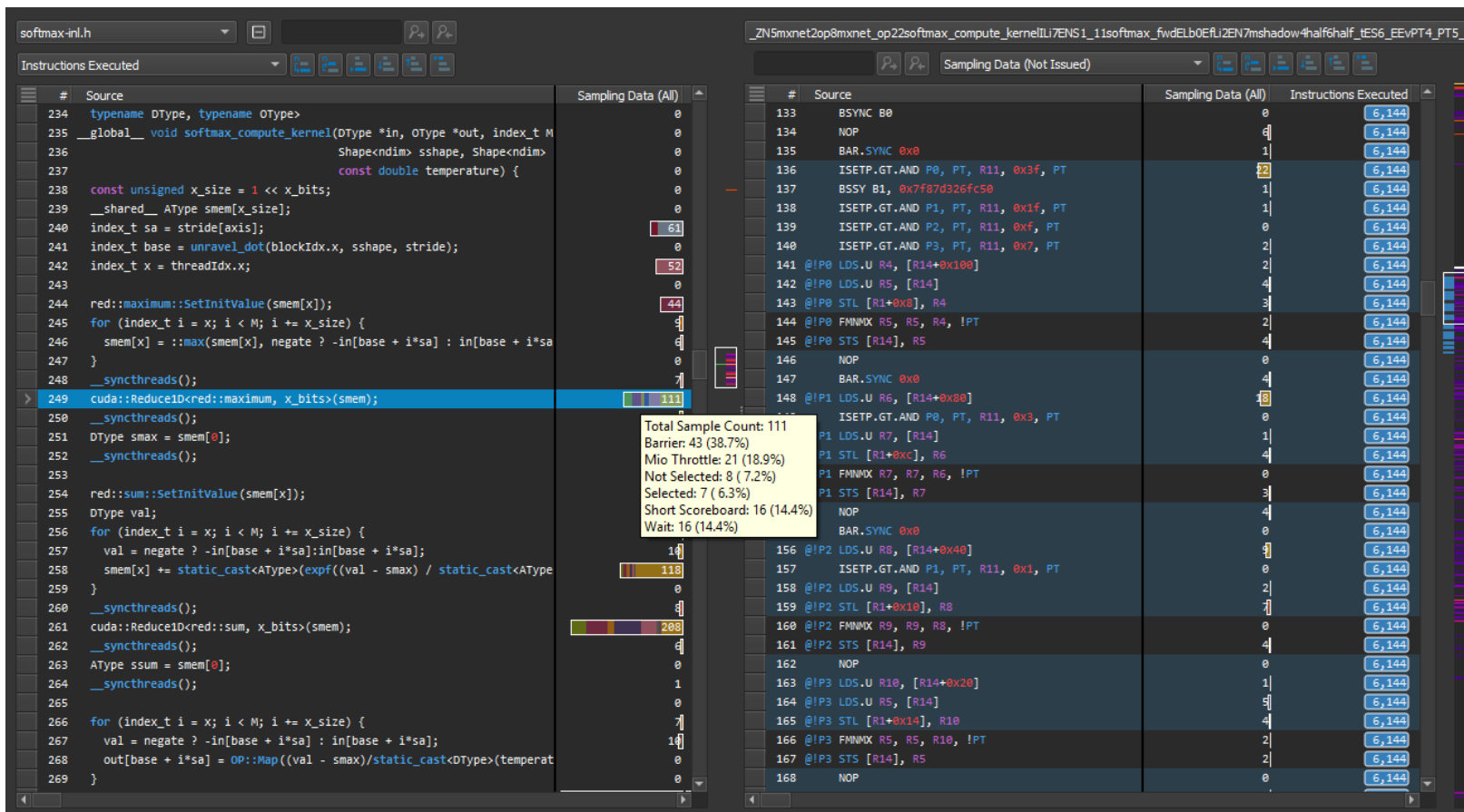
Nsight Compute



Comparison of results directly within the tool with “Baselines”

Supported across kernels, reports, and GPU architectures

Nsight Compute



Source/PTX/SASS
analysis and
correlation

Source metrics per
instruction and
aggregated (e.g. PC
sampling data)

Metric heatmap

Nsight Compute

```
==PROF== Disconnected from process 8792
[8792] CuBlackscholes.exe@127.0.0.1
GPUBlackScholesCallPut(int, float*, float*, float*, float*, float*), Block Size 256,
Section: GPU Speed Of Light
```

Metric Name	Metric Unit	Minimum	Maximum	Average
dram__frequency	Ghz	6.004059	6.004059	6.004059
fbpa__sol_pct	%	70.191350	70.191350	70.191350
gpc__elapsed_cycles_max	cycle	751140.000000	751140.000000	751140.000000
gpc__frequency	Ghz	1.287365	1.287365	1.287365
gpu__compute_memory_sol_pct	%	70.191350	70.191350	70.191350
gpu__time_duration	usecond	583.456000	583.456000	583.456000
ltc__sol_pct	%	24.488190	24.488190	24.488190
sm__elapsed_cycles_avg	cycle	751121.000000	751121.000000	751121.000000
sm__sol_pct	%	69.036830	69.036830	69.036830
tex__sol_pct	%	20.449435	20.449435	20.449435

Full command line interface (CLI) for data collection and analysis

On your workstation

Support for remote profiling across machines, platforms (Linux, Windows, ...) in UI and CLI

```
GPUBlackScholesCallPut(int, float*, float*, float*, float*, float*), 2019-Aug-12 14:44:50, Context 1, Stream 7
Section: GPU Speed Of Light
```

Metric Name	Metric Unit	Value
Memory Frequency	Ghz	6.00
SOL FB	%	71.55
Elapsed Cycles	cycle	749,683
SM Frequency	Ghz	1.29
Memory [%]	%	71.55
Duration	usecond	580.93
SOL L2	%	24.53
SM Active Cycles	cycle	749,656
SM [%]	%	69.17
SOL TEX	%	20.49

OK Compute and Memory are well-balanced: To reduce runtime, both computation and memory traffic must be reduced. Check both the `Compute Workload Analysis` and `Memory Workload Analysis` report sections.

KERNEL PROFILES WITH NSIGHT COMPUTE

```
$ ncu -k mykernel ./myapp.exe
```

(Without the -k option, Nsight Compute will profile everything and take a long time!)

The Nsight Compute UI can also be used for interactive kernel profiling



MORE RESOURCES

[OLCF Nsight Systems Webinar](#)

[OLCF Nsight Compute Webinar](#)

[Nsight Compute at GTC 2019](#)

[Nsight Compute Documentation](#) and [Nsight Systems Documentation](#)